



What is the household composition of families in the family justice system? Linking data to fill the evidence gaps

Guidance note

February 2021

About this document

This guidance note provides the methodological detail behind the analysis for 'What is the household composition of families in the family justice system? Linking data to fill the evidence gaps' an online article by Nuffield Family Justice Observatory (Nuffield FJO). The article is available at:

<https://www.nuffieldfjo.org.uk/news/data-linkage-household-composition-family-justice-system>

The analysis was carried out by Rhodri Johnson and Lucy Griffiths (Swansea University) as part of the Family Justice Data Partnership—a collaboration between Lancaster University and Swansea University, funded by the Nuffield Family Justice Observatory.

Copyright © Nuffield Family Justice Observatory 2021
28 Bedford Square, London WC1B 3JS T: 020 7631 0566
Registered charity 206601
nuffieldfjo.org.uk | @NuffieldFJO www.nuffieldfoundation.org | @NuffieldFound

Contents

1. Introduction	3
2. Data sources	3
3. Study group selection process.....	3
Public and private law study group selection	3
Comparison group selection	4
4. Measures	4
Resident characteristics	4
Resident generational and household classification.....	4
5. Limitations	5
References	5

1. Introduction

This guidance note provides the methodological detail behind the analysis for 'What is the household composition of families in the family justice system? Linking data to fill the evidence gaps' an [online article](#) by Nuffield Family Justice Observatory (Nuffield FJO).

The analysis was carried out by Rhodri Johnson and Lucy Griffiths (Swansea University) as part of the Family Justice Data Partnership—a collaboration between Lancaster University and Swansea University, funded by the Nuffield Family Justice Observatory.

2. Data sources

The SAIL (Secure Anonymised Information Linkage) Databank contains extensive anonymised health and administrative data about the population of Wales, accessible via a secure data sharing platform, all underpinned by an innovative and proportionate information governance model (Ford et al. 2009; Lyons et al. 2009; Jones et al. 2019).

Within SAIL, records for individuals can be linked using anonymised linking fields (ALF), which can be linked to a property or a residential anonymised linking field (RALF) (Rodgers et al. 2009). The Welsh Demographic Service (WDS) data contains GP address registration information of Welsh residents, demographic data, and dates at which people are resident within properties in Wales.

This analysis used this residential information, linked to data on public and private law proceedings as captured in Cafcass Cymru records (Johnson et al. 2020), to explore which residents lived together in households, to infer relationships between residents based on age differences, and to form measures of household composition.

3. Study group selection process

Three study groups were created:

- households involved in public law proceedings (n=3,666)
- households involved in private law proceedings (16,592)
- a comparison group (n=311,729), representing households from the general population of Wales.

Public and private law study group selection

We selected the youngest child (aged under 18) identified as a 'subject' in Cafcass Cymru applications between 2011 and 2019 (inclusive) where the child had an ALF and a valid birth date; the child's first application date was used as the census date. We linked the child's ALF to WDS data to obtain the RALF to which they were registered at this date. All other ALFs registered to the property were classed as co-residents. The case definition of law type was used to classify law type.

Comparison group selection

We created a comparison group by selecting RALFs with at least one child (aged under 18) registered at a census date (30 June 2015), where the child was not involved in care proceedings, or registered to the same RALF as any of the study group. All other ALFs registered to the property were classed as co-residents.

Following analysis of household size (total number of ALFs), it was decided to retain only households with between two and nine people in them.

4. Measures

Resident characteristics

Characteristics included the percentage of residents by gender and tenancy duration calculated as the number of years from initial registration in a property to the census date (used as a proxy measure for housing security). Analysis was carried out to estimate the number of siblings per household, and housing composition, segmented by age of the youngest child according to three age categories:

- infants (under 1 year)
- young children (1–9 years)
- adolescents (10–17 years).

Resident generational and household classification

Each household was classified using an amended version of the 'Age to Youngest' (AtY) method (Johnson et al. forthcoming) where each resident in a household was compared to the youngest resident to calculate a relative age difference; this method was used to separate household members into one of three estimated generations (Table 1). The resulting household generational classifications were combined to create household composition classification estimates. For example, a household with four residents aged 6, 8, 43, and 45 would equate to 2-2-0, indicating two people in each of the first two generational classifications only.

Table 1: Generation classification

Classification	Generation classification	Age rule
Child/sibling	1	Within 0–16 years of youngest resident
Parent	2	Within 17–45 years of youngest resident
Grandparent	3	More than 45 years older than youngest resident

Household classification was used to model family/parental composition, aiming to understand how many parents were present in households, or how many households had residents of grandparent age but no parent-aged adult (Table 2).

Table 2: Household composition classifications

Stage	Rule description	Household composition classification
1	One person of generation 2	Lone parents with child(ren)
2	Two or more person of generation 2	Two parents with child(ren)
3	Zero persons of generation 2, and one or more persons of generation 3	One or more grandparent with child(ren)
4	Zero persons of generation 2 or 3*	No adult classified with child(ren)

*Excluded in blog charts.

5. Limitations

It should be noted that household measures are estimates at a group level and, undoubtedly, some households will be misclassified. The terminology used relates to family members (siblings, parents, and grandparents), however, relationships were *inferred* based on co-residence and age difference.

For this preliminary analysis, the comparison group was not matched on age, deprivation, or other factors. Comparisons should therefore be interpreted with caution.

Studies based on administrative data are necessarily limited by the scope and quality of available data, which is collected primarily for organisational rather than research purposes. Limitations of Cafcass Cymru data have been previously described (Johnson et al. 2020).

References

- Alrouh, B. et al. (2019). *Born into care: Newborns and infants in care proceedings in Wales*. London: Nuffield Family Justice Observatory.
- Ford, D. et al. (2009). The SAIL Databank: building a national architecture for e-health research and evaluation. *BMC Health Services Research*, 9(1), p. 157.
- Johnson, R. D. et al. (2020). Deriving household composition using population-scale Electronic Health Record data – a reproducible methodology. *Plos One*, in press pre-print: <https://doi.org/10.31235/osf.io/kctnz>.
- Johnson, R. D. et al. (2020). Data resource: population level family justice administrative data with opportunities for data linkage. *International Journal of Population Data Science*, 5(1).
- Jones, K. H. et al. (2019). A profile of the SAIL Databank on the UK Secure Research Platform. *International Journal of Population Data Science*, 4(2).
- Lyons, R. A. et al. (2009). The SAIL Databank: linking multiple health and social care datasets. *BMC Medical Informatics and Decision Making*. BioMed Central, 9(1), p. 3.
- Rodgers, S. E. et al. (2009). Residential anonymous linking fields (RALFs): A novel information infrastructure to study the interaction between the environment and individuals' health. *Journal of Public Health*, 31(4), pp. 582–588.